PART 2

UNDERSTANDING THE EVIDENCE

CHAPTER 2

Diagnostic Accuracy of Physical Findings

KEY TEACHING POINTS

- Likelihood ratios (LRs) are nothing more than diagnostic weights, numbers that quickly convey to clinicians how much a physical sign argues for or against disease.
- LRs have possible values between 0 and ∞. Values greater than 1 *increase* the probability of disease. (The greater the value of the LR, the greater the increase in probability.) LRs less than 1 *decrease* the probability of disease. (The closer the number is to zero, the more the probability of disease decreases.) LRs that equal 1 do not change the probability of disease at all.
- LRs of 2, 5, and 10 increase the probability of disease about 15%, 30%, and 45%, respectively (in absolute terms). LRs of 0.5, 0.2, and 0.1 (i.e., the reciprocals of 2, 5, and 10) decrease probability 15%, 30%, and 45%, respectively.
- Tables comparing LRs of different physical signs quickly inform clinicians about which findings have the greatest diagnostic value.

I. INTRODUCTION

If a physical sign characteristic of a suspected diagnosis is present (i.e., **positive finding**), that diagnosis becomes more likely; if the characteristic finding is absent (i.e., **negative finding**), the suspected diagnosis becomes less likely. How much these positive and negative results modify probability, however, is distinct for each physical sign. Some findings, when positive, increase probability significantly, but they change it little when negative. Other signs are more useful if they are absent, because a negative finding practically excludes disease, although a positive one changes probability very little.

TABLE 2.1 Pre-Test Probability

Setting (Reference)	Diagnosis	Probability (%)
Acute abdominal pain[1-3]	Small bowel obstruction	4
Ankle injury[4,5]	Ankle fracture	10-14
Cough and fever[6]	Pneumonia	12-30
Acute calf pain or swelling[7-15]	Proximal deep vein thrombosis	13-43
Pleuritic chest pain, dyspnea, or hemoptysis[16-19]	Pulmonary embolism	9-43
Diabetic foot ulcer[20-22]	Osteomyelitis	52-68

Much of this book consists of tables that specifically describe how positive or negative findings change the probability of disease, a property called **diagnostic accuracy**. Understanding these tables first requires a review of four concepts: pre-test probability, sensitivity, specificity, and LRs.

II. PRE-TEST PROBABILITY

Pre-test probability is the probability of disease (i.e., prevalence) before application of the results of a physical finding. Pre-test probability is the starting point for all clinical decisions. For example, the clinician may know that a certain physical finding increases the probability of disease 40%, but this information alone is unhelpful unless the clinician also knows the starting point: if the pre-test probability for the particular diagnosis was 50%, the finding is diagnostic (i.e., post-test probability 50% + 40% = 90%); if the pre-test probability was only 10%, the finding is less helpful, because the probability of disease is still akin to a coin toss (i.e., post-test probability 10% + 40% = 50%).

Published estimates of disease prevalence, given a particular clinical setting, are summarized in the Appendix for all clinical problems discussed in this book. (These estimates derive from clinical studies reviewed in the EBM Boxes.) Table 2.1 provides a small sample of these pre-test probabilities. Even so, clinicians must adjust these estimates with information from their own practice. For example, large studies based in emergency departments show that 12% to 35% of patients presenting with cough and fever have pneumonia (see Table 2.1). The probability of pneumonia, however, is certainly lower in patients presenting with cough and fever to an office-based practice, and it may be higher if cough and fever develops in patients with cancer or human immunodeficiency virus (HIV) infection. In fact, because the best estimate of pre-test probability incorporates information from the clinician's own practice—how specific underlying diseases, risks, and exposures make disease more or less likely—the practice of evidence-based medicine is never "cookbook" medicine, but instead consists of decisions based on the unique characteristics of the patients the clinician sees.

III. SENSITIVITY AND SPECIFICITY

A. DEFINITIONS

The terms *sensitivity* and *specificity* are used to describe the discriminatory power of physical signs. **Sensitivity** is the proportion of patients *with* the diagnosis who *have* the physical sign (i.e., have the *positive* result). **Specificity** is the proportion of patients *without* the diagnosis who *lack* the physical sign (i.e., have the *negative* result).

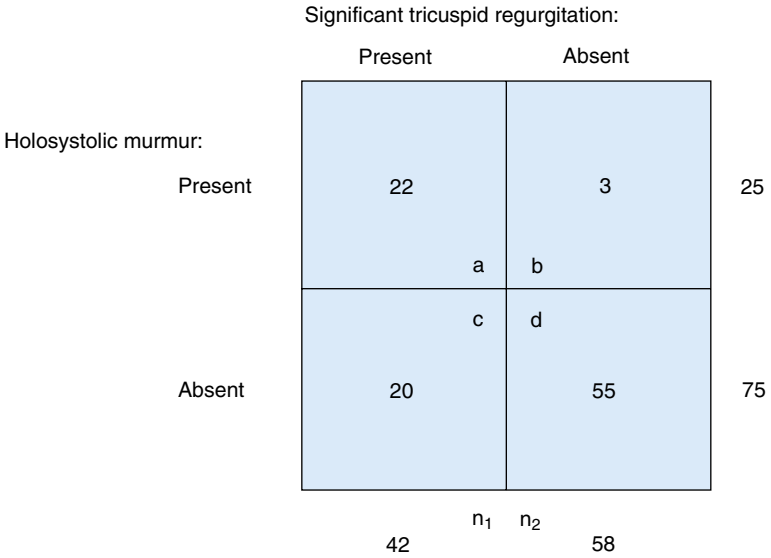

FIG. 2.1 2 × 2 TABLE. The total number of patients with disease (tricuspid regurgitation in this example) is the sum of the first column, or $n_1 = a + c$. The total number of patients without disease is the sum of the second column, or $n_2 = b + d$. The *sensitivity* of a physical finding (holosystolic murmur at the left lower sternal edge, in this example) is the proportion of patients with disease who have the finding [i.e., $a/(a + c)$, or a/n_1]. The *specificity* of a physical finding is the proportion of patients without disease who lack the finding [i.e., $d/(b + d)$, or d/n_1]. The *positive LR* is the proportion of patients with disease who have a positive finding (a/n_1) divided by the proportion of patients without disease who have a positive finding (b/n_2), or sensitivity/(1 − specificity). The *negative LR* is the proportion of patients with disease who lack the finding (c/n_1) divided by the proportion of patients without disease who lack the finding (d/n_1), or (1 − sensitivity)/specificity. In this example, the sensitivity is 0.52 (22/42), the specificity is 0.95 (55/58), the positive LR is 10.1 [(22/42)/(3/58)], and the negative LR is 0.5 [(20/42)/(55/58)].

The calculation of sensitivity and specificity requires the construction of a 2 × 2 table (Fig. 2.1) that has two columns (one for "diagnosis present" and another for "diagnosis absent") and two rows (one for "physical sign present" and another for "physical sign absent"). These rows and columns create four boxes: one for the "true positives" (cell a, sign and diagnosis present), one for the "false positives" (cell b, sign present but disease absent), one for the "false negatives" (cell c, sign absent but disease present), and one for the "true negatives" (cell d, sign and disease absent).

Fig. 2.1 presents data from a hypothetical study of 100 patients presenting with pulmonary hypertension. The clinician knows that tricuspid regurgitation is a complication of pulmonary hypertension and wonders how accurately a single physical sign—the presence of a holosystolic murmur at the left lower sternal border—detects this complication.* In this study, 42 patients have significant tricuspid regurgitation (the sum of column 1) and 58 patients do not (the sum of column 2). The **sensitivity** of the holosystolic murmur is the proportion of patients with disease

*The numbers used in this example are very close to those given in reference 23. See also Chapter 46.

(i.e., tricuspid regurgitation, 42 patients) who have the characteristic murmur (i.e., the *positive* result, 22 patients), which is 22/42 = 0.52 or 52%. The **specificity** of the holosystolic murmur is the proportion of patients *without* disease (i.e., no tricuspid regurgitation, 58 patients) who *lack* the murmur (i.e., the *negative* result, 55 patients), which is 55/58 = 0.95 or 95%.

To recall how to calculate sensitivity and specificity, Sackett and others have suggested helpful mnemonics: Sensitivity is represented as "PID" for "positivity in disease" (an abbreviation normally associated with "pelvic inflammatory disease"), and specificity is represented as "NIH" for "negativity in health" (an abbreviation normally associated with the "National Institutes of Health").[24,25]

B. USING SENSITIVITY AND SPECIFICITY TO DETERMINE PROBABILITY OF DISEASE

The completed 2 × 2 table can be used to determine the accuracy of the holosystolic murmur, which is how well its presence or absence discriminates between those with tricuspid regurgitation and those without it. In Fig. 2.1, the first row includes all 25 patients with the murmur (i.e., the positive results). Of these 25 patients, 22 have tricuspid regurgitation; therefore the probability of tricuspid regurgitation if the murmur is present (*positive* finding) is 22/25 or 88% (i.e., the "post-test probability" if the murmur is present). The second row includes all 75 patients without the murmur. Of these 75 patients, 20 have tricuspid regurgitation; therefore, the post-test probability of tricuspid regurgitation if the murmur is absent (i.e., *negative* finding) is 20/75 or 27%.

In this example, the pre-test probability of tricuspid regurgitation is 42%. The presence of the murmur (positive result) increases the probability of disease considerably more (i.e., 46%, from 42% to 88%) than the absence of the murmur (negative result) decreases it (i.e., 15%, from 42% to 27%). This illustrates an important property of physical signs with a high specificity: when present, physical signs with *high specificity* greatly *increase* the probability of disease. A corollary to this applies to findings with high sensitivity: when *absent*, physical signs with a high *sensitivity* greatly *decrease* the probability of disease. The holosystolic murmur has a high specificity (95%) but only a meager sensitivity (52%), meaning that, at the bedside, a positive result (the presence of a murmur) has greater diagnostic importance than a negative result (the absence of a murmur). The presence of the characteristic murmur argues compellingly for tricuspid regurgitation, but its absence is less helpful, simply because many patients with significant regurgitation lack the characteristic murmur.

Sackett and others have suggested mnemonics for these characteristics as well: "SpPin" (i.e., a *Sp*ecific test, when *P*ositive, rules *in* disease) and "SnNout" (i.e., a *Sn*sitive test, when *N*egative, rules *out* disease).[25]

IV. LIKELIHOOD RATIOS

LRs, like sensitivity and specificity, describe the discriminatory power of physical signs. Although they have many advantages, the most important is how simply and quickly they can be used to estimate post-test probability.

A. DEFINITION

The LR of a physical sign is the proportion of patients *with* disease who have a particular finding divided by the proportion of patients *without* disease who also have the same finding.

$$\text{LR} = \frac{\text{Probability of finding in patients } \textit{with} \text{ disease}}{\text{Probability of the same finding in patients } \textit{without} \text{ disease}}$$

The adjectives *positive* or *negative* indicate whether that LR refers to the presence of the physical sign (i.e. positive result) or to the absence of the physical sign (i.e., the negative result).

A **positive LR**, therefore, is the proportion of patients *with* disease who *have* a physical sign divided by the proportion of patients *without* disease who also *have* the same sign. The numerator of this equation—the proportion of patients with disease who have the physical sign—is the sign's sensitivity. The denominator—the proportion of patients without disease who have the sign—is the complement of specificity, or (1 − specificity). Therefore,

$$\text{Positive LR} = \frac{(\text{sens})}{(1 - \text{spec})}$$

In our hypothetical study (see Fig. 2.1), the proportion of patients with tricuspid regurgitation who have the murmur is 22/42, or 52.4% (i.e., the finding's sensitivity), and the proportion of patients without tricuspid regurgitation who also have the murmur is 3/58, or 5.2% (i.e., 1 − specificity). The ratio of these proportions [i.e., (sensitivity)/(1 − specificity)] is 10.1, which is the positive LR for a holosystolic murmur at the lower sternal border. This number indicates that patients *with* tricuspid regurgitation are 10.1 times more likely to have the holosystolic murmur than those *without* tricuspid regurgitation.

Similarly, the **negative LR** is the proportion of patients *with* disease *lacking* a physical sign divided by the proportion of patients *without* disease also *lacking* the sign. The numerator of this equation—the proportion of patients with disease *lacking* the finding—is the complement of sensitivity, or (1 − sensitivity). The denominator of the equation—the proportion of patients without disease *lacking* the finding—is the specificity. Therefore,

$$\text{Negative LR} = \frac{(1 - \text{sens})}{(\text{spec})}$$

In our hypothetical study, the proportion of patients with tricuspid regurgitation lacking the murmur is 20/42, or 47.6% (i.e., 1 − sensitivity), and the proportion of patients without tricuspid regurgitation lacking the murmur is 55/58, or 94.8% (i.e., the specificity). The ratio of these proportions [i.e., (1 − sensitivity)/(specificity)] is 0.5, which is the negative LR for the holosystolic murmur. This number indicates that patients *with* tricuspid regurgitation are 0.5 times less likely to lack the murmur than those *without* tricuspid regurgitation. (The inverse statement is less confusing: patients *without* tricuspid regurgitation are 2 times more likely to lack a murmur than those *with* tricuspid regurgitation.)

Although these formulas are difficult to recall, the interpretation of LRs is straightforward. Findings with LRs greater than 1 increase the probability of disease; the greater the LR, the more compelling the argument *for* disease. Findings whose LRs lie between 0 and 1 decrease the probability of disease; the closer the LR is to zero, the more convincing the finding argues *against* disease. Findings whose LRs equal 1 lack diagnostic value because they do not change probability at all. "Positive LR" describes how probability changes when the finding is *present*. "Negative LR" describes how probability changes when the finding is *absent*.

LRs, therefore, are nothing more than diagnostic weights, whose possible values range from 0 (i.e., excluding disease) to infinity (i.e., pathognomic for disease; Fig. 2.2).

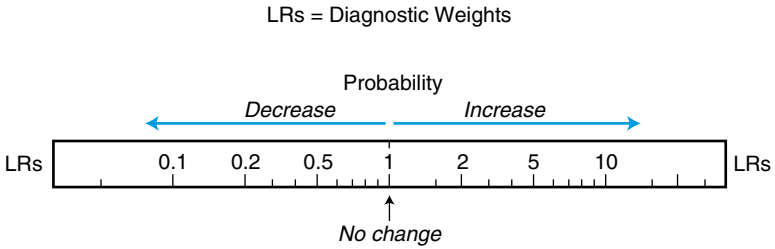

FIG. 2.2 LIKELIHOOD RATIOS AS DIAGNOSTIC WEIGHTS. The relationship between a specific physical sign and a specific disease is described by a unique number—its likelihood ratio—which is nothing more than a diagnostic weight describing how much that sign argues for or against that specific disease. The possible values of LRs range from zero to infinity (∞). Findings with LRs greater than 1 argue *for* the specific disease (the greater the value of the LR, the more the probability of disease increases). Findings with LRs less than 1 argue *against* the disease (the closer the number is to zero, the more the probability of disease decreases). LRs that equal 1 do not change probability of disease at all.

B. USING LIKELIHOOD RATIOS TO DETERMINE PROBABILITY

The clinician can use the LR of a physical finding to estimate probability of disease in three ways: (1) by using graphs or other easy-to-use nomograms,[26,27] (2) by using bedside approximations, or (3) by using formulas.

1. USING GRAPHS

A. PARTS OF THE GRAPH

Fig. 2.3 is an easy-to-use graph that illustrates the relationship between pre-test probability (*x*-axis) and post-test probability (*y*-axis), given the finding's LR. The straight line bisecting the graph into an upper left half and a lower right half indicates an LR of 1, which has no discriminatory value because, for findings with this LR, post-test probability always equals pre-test probability. Physical findings that argue *for* disease (i.e., LRs >1) appear in the upper left half of the graph; the larger the value of the LR, the more the curve approaches the upper left corner. Physical findings that argue *against* disease (i.e., LRs <1) appear in the lower right half of the graph: the closer the LR is to zero, the more the curve approaches the lower right corner.

In Fig. 2.3, the three depicted curves with LRs greater than 1 (i.e., LR = 2, 5, and 10) are mirror images of the three curves with LRs less than 1 (i.e., LR = 0.5, 0.2, and 0.1). (This assumes the "mirror" is the line LR = 1.) This symmetry indicates that findings with an LR of 10 argue as much *for* disease as those with an LR = 0.1 argue *against* disease (although this is true only for the intermediate pre-test probabilities). Similarly, LR = 5 argues as much for disease as LR = 0.2 argues against it, and LR = 2 mirrors LR = 0.5. Keeping these companion curves in mind will help the clinician interpret the LRs throughout this book.[†]

[†]These companion pairs are easy to recall because they are inversely related: the inverse of 10 is 1/10 = 0.1; the inverse of 5 is 1/5 = 0.2; the inverse of 2 is 1/2 = 0.5.

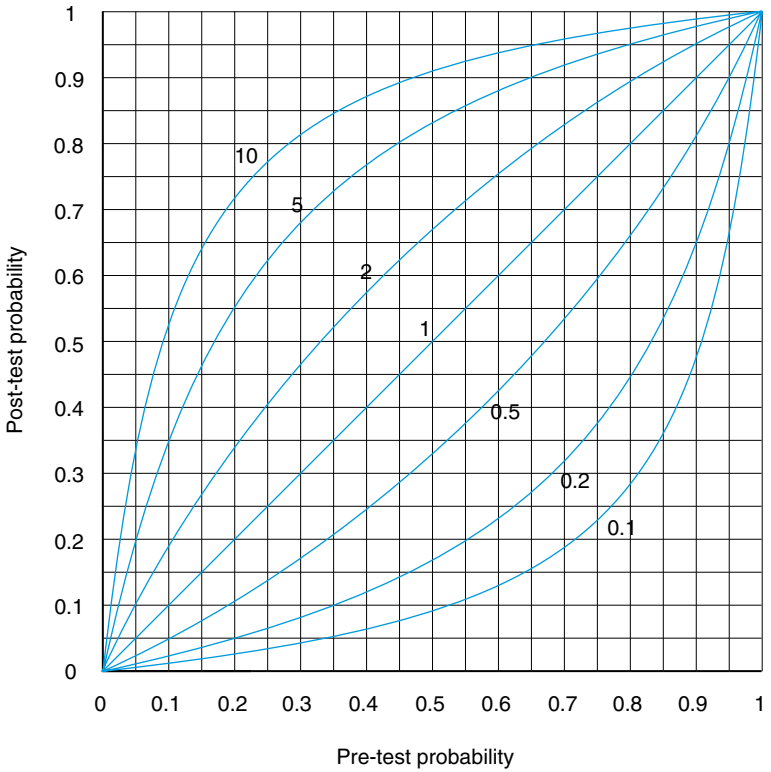

FIG. 2.3 PROBABILITY AND LIKELIHOOD RATIOS. The curves describe how pre-test probability (*x*-axis) relates to post-test probability (*y*-axis), given the likelihood ratio (LR) for the physical finding. Only the curves for seven likelihood ratios are depicted (from LR = 0.1 to LR = 10).

If a finding has an LR other than one of these depicted seven curves, its position can be estimated with little loss in accuracy. For example, the curve for LR = 4 lies between LR = 5 and LR = 2, though it is closer to LR= 5 than it is to LR = 2.

B. USING THE GRAPH TO DETERMINE PROBABILITY

To use this graph, the clinician identifies on the *x*-axis the patient's pre-test probability, derived from published estimates and clinical experience, and extends a line upward from that point to meet the LR curve for the physical finding. The clinician then extends a horizontal line from this point to the *y*-axis to identify post-test probability.

Fig. 2.4 depicts this process for the lower sternal holosystolic murmur and tricuspid regurgitation. The pre-test probability of tricuspid regurgitation is 42%. If the characteristic murmur is present (positive LR = 10), a line is drawn upward from 0.42 on the *x*-axis to the LR = 10 curve; from this point, a horizontal line is drawn to the *y*-axis to find the post-test probability (88%). If the murmur is absent (negative LR = 0.5), the post-test probability is the *y*-value where the vertical line intersects the LR = 0.5 curve (i.e., a post-test probability of 27%).

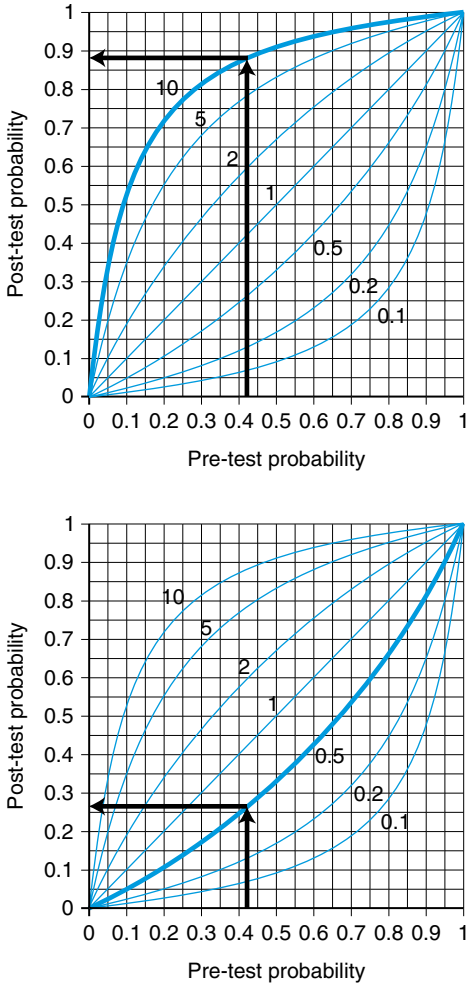

FIG. 2.4 PROBABILITY AND LIKELIHOOD RATIOS: PATIENTS WITH PULMONARY HYPERTENSION. In our hypothetical clinician's practice, 42% of patients with pulmonary hypertension have the complication of tricuspid regurgitation (i.e., pre-test probability is 42%). To use the curves, the clinician finds 0.42 on the *x*-axis and extends a line upward. The post-test probability of tricuspid regurgitation is read off the *y*-axis where the vertical line intersects the curve of the appropriate LR. The probability of tricuspid regurgitation if a holosystolic murmur is present at the left lower sternal edge (LR = 10.1) is 88%; the probability if the finding is absent (LR = 0.5) is 27%.

These curves illustrate an additional important point: physical signs are diagnostically most useful when they are applied to patients who have pre-test probabilities in the intermediate range (i.e., 20% to 80%), because in this range, the different LR curves diverge the most from the LR = 1 curve (thus significantly increasing or decreasing probability). If instead the pre-test probability is already

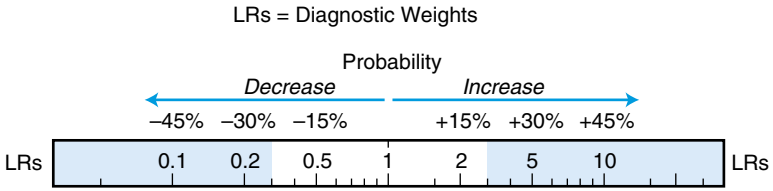

FIG. 2.5 APPROXIMATING PROBABILITY. Clinicians can estimate changes in probability by recalling the LRs 2, 5, and 10 and the first 3 multiples of 15 (i.e., 15, 30, and 45). A finding whose LR is 2 increases probability about 15%, one of 5 increases it 30%, and one of 10 increases it 45% (these changes are *absolute* increases in probability). LRs whose values are 0.5, 0.2, and 0.1 (i.e., the reciprocals of 2, 5, and 10) decrease probability 15%, 30%, and 45%, respectively. Throughout this book, LRs with values ≥3 or ≤0.3 (represented by the shaded part of the diagnostic weight "ruler") are presented in boldface type to indicate those physical findings that change probability sufficiently to be clinically meaningful (i.e., they increase or decrease probability at least 20% to 25%).

very low or very high, all the LR curves cluster close to the line LR = 1 curve in either the bottom left or upper right corners, thus with only a relatively small impact on probability.

2. APPROXIMATING PROBABILITY

The clinician can avoid using graphs and instead approximate post-test probability by remembering the following two points: First, the companion LR curves in Fig. 2.3 are LR = 2 and LR = 0.5, LR = 5 and LR = 0.2, and LR = 10 and LR = 0.1. Second, the first three multiples of 15 are 15, 30, and 45. Using this rule, the LRs of 2, 5, and 10 increase probability about 15%, 30%, and 45%, respectively (Fig. 2.5). The LRs of 0.5, 0.2, and 0.1 decrease probability by about 15%, 30%, and 45%, respectively.[28] These estimates are accurate to within 5% to 10% of the actual value, as long as the clinician rounds estimates over 100 to an even 100% and estimates below zero to an even 0%.

Therefore, in our hypothetical patient with pulmonary hypertension, the finding of a holosystolic murmur (LR = 10) increases the probability of tricuspid regurgitation from 42% to 87% (i.e., 42% + 45% = 87%, which is only 1% lower than the actual value). The absence of the murmur (LR = 0.5) decreases the probability of tricuspid regurgitation from 42% to 27% (i.e., 42% − 15% = 27%, which is identical to actual value).

Table 2.2 summarizes similar bedside estimates for all LRs between 0.1 and 10.0.

3. CALCULATING PROBABILITY

Post-test probability can also be calculated by first converting pre-test probability (P_{pre}) into pre-test odds (O_{pre}):

$$O_{pre} = \frac{P_{pre}}{(1 - P_{pre})}$$

The pre-test odds (O_{pre}) is multiplied by the LR of the physical sign to determine the post-test odds (O_{post}):

$$O_{post} = O_{pre} \times LR$$

TABLE 2.2 Likelihood Ratios and Bedside Estimates

Likelihood Ratio	Approximate Change in Probability*
0.1	−45%
0.2	−30%
0.3	−25%
0.4	−20%
0.5	−15%
1	No change
2	+15%
3	+20%
4	+25%
5	+30%
6	+35%
7	
8	+40%
9	
10	+45%

*These changes describe *absolute* increases or decreases in probability. For example, a patient with a pre-test probability of 20% and physical findings whose LR is 5 would have a post-test probability of 20% + 30% = 50%. The text describes how to easily recall these estimates.
Based upon reference 28.

The post-test odds (O_{post}) converts back to post-test probability (P_{post}), using

$$P_{post} = \frac{O_{post}}{(1 + O_{post})}$$

Therefore, in our hypothetical example of the patients with pulmonary hypertension, the pre-test odds for tricuspid regurgitation would be [(0.42)/(1 − 0.42)], or 0.72. If the murmur is present (LR = 10), the post-test odds would be [0.72 × 10], or 7.2, which translates to a post-test probability of [(7.2)/(1 + 7.2)], or 0.88 (i.e., 88%). If the murmur wave is absent (LR = 0.5), the post-test odds would be [0.72 × 0.5], or 0.36, which translates to a post-test probability of [(0.36)/(1 + 0.36)], or 0.27 (i.e., 27%).

Clinical medicine, however, is rarely as precise as these calculations suggest, and for most decisions at bedside, the approximations described earlier are more than adequate.

C. ADVANTAGES OF LIKELIHOOD RATIOS

1. SIMPLICITY

In a single number, the LR conveys to clinicians how convincingly a physical sign argues for or against disease. If the LR of a finding is large, disease is likely, and if the LR of a finding is close to zero, disease is doubtful. This advantage allows clinicians to quickly compare different diagnostic strategies and thus refine clinical judgment.[28]

2. ACCURACY

Using LRs to describe diagnostic accuracy is superior to describing it in terms of sensitivity and specificity, because the previously explained mnemonics, SpPin and SnNout, are sometimes misleading. For example, according to the mnemonic SpPin, a finding with a specificity of 95% should argue conclusively for disease, but

TABLE 2.3 Breath Sound Intensity and Chronic Airflow Limitation

Breath Sound Score	Likelihood Ratio
9 or less	10.2
10-12	3.6
13-15	NS
>15	0.1

NS, not significant.
Based upon references 29 and 30.

it does so only if the positive LR for the finding is a high number. If the finding's sensitivity is 60%, the positive LR is 12 and the finding argues convincingly for disease (i.e., consistent with the SpPin mnemonic); if the finding's sensitivity is only 10%, however, the positive LR is 2 and the post-test probability changes only slightly (i.e., inconsistent with SpPin mnemonic). Similarly, a highly sensitive finding argues convincingly against disease when absent (i.e., SnNout) only when its calculated negative LR is close to zero.

3. LEVELS OF FINDINGS

Another advantage of LRs is that a physical sign measured on an ordinal scale (e.g., 0, 1+, 2+, 3+) or continuous scale (e.g., blood pressure) can be categorized into different levels to determine the LR for each level, thereby increasing the accuracy of the finding. Other examples include continuous findings such as heart rate, respiratory rate, temperature, and percussed span of the liver, and ordinal findings such as intensity of murmurs and degree of edema.

For example, in patients with chronic obstructive lung disease (i.e., emphysema, chronic bronchitis), breath sounds are typically faint. If the clinician grades the intensity of breath sounds on a scale from 0 (absent) to 24 (very loud), based on the methods discussed in Chapter 30,[29,30] he or she can classify the patient's breath sounds into one of four groups: scores of 9 or less (very faint), 10 to 12, 13 to 15, or greater than 15 (loud). Each category then has its own LR (Table 2.3): scores of 9 or less significantly increase the probability of obstructive disease (LR = 10.2), whereas scores greater than 15 significantly decrease it (LR = 0.1). Scores from 10 to 12 argue somewhat for disease (LR = 3.6), and scores from 13 to 15 provide no diagnostic information (LR not significantly different from 1). If the clinician instead identifies breath sounds as simply "faint" or "normal/increased" (i.e., the traditional positive or negative finding), the finding may still discriminate between patients with and without obstructive disease, but it misses the point that the discriminatory power of the sign resides mostly with scores less than 10 and greater than 15.

When findings are categorized into levels, the term *specificity* becomes meaningless. For example, the specificity of a breath sound score of 13 to 15 is 80%, which means that 80% of patients without chronic airflow limitation have values other than 13 to 15, though the "80%" does not convey whether most of these other values are greater than 15 or less than 13. Similarly, when findings are put in more than two categories, the LR descriptor *negative* is no longer necessary, because all LRs are *positive* for their respective category.

4. COMBINING FINDINGS

A final advantage of LRs is that clinicians can use them to combine findings, which is particularly important for those physical signs with positive LRs around 2 or negative LRs around 0.5, signs that by themselves have little effect on probability but

when combined have significant effects on probability. Individual LRs can be combined—however, only if the findings are "independent."

A. INDEPENDENCE OF FINDINGS

Independence means that the LR for the second finding does not change once the clinician determines whether the first finding is present or absent. For some select diagnostic problems, investigators have identified which findings are independent of each other. These findings appear as components of "diagnostic scoring schemes" in the tables throughout this book (e.g., Wells score for deep venous thrombosis). For most physical findings, however, very little information is available about independence, and the clinician must judge whether combining findings is appropriate.

One important indication is that most independent findings have unique pathophysiology. For example, when considering pneumonia in patients with cough and fever, the clinician could combine the findings of abnormal mental status and diminished breath sounds, using the individual LR of each finding because abnormal mental status and diminished breath sounds probably have separate pathophysiology. Similarly, when considering heart failure in patients with dyspnea, the clinician could combine the findings of elevated neck veins and the third heart sound because these findings also have different pathophysiology.

Examples of findings whose individual LRs should *not* be combined (because the findings share the same pathophysiology) are flank dullness and shifting dullness in the diagnosis of ascites (both depend on intra-abdominal contents dampening the vibrations of the abdominal wall during percussion), neck stiffness and the Kernig sign in the diagnosis of meningitis (both are caused by meningeal irritation), and edema and elevated neck veins in the diagnosis of heart failure (both depend on elevated right atrial pressure).

Until more information is available, the safest policy for the clinician to follow when combining LRs of individual findings is to combine no more than three findings, all of which have distinct pathophysiology.

B. HOW TO COMBINE FINDINGS

The clinician can use any of the methods previously described to combine findings, simply by making the post-test probability from the first finding the pre-test probability for the second finding. For example, a hypothetical patient with acute fever and cough has two positive findings that we believe have separate pathophysiology and therefore are independent: abnormal mental status (LR = 1.9 for pneumonia) and diminished breath sounds (LR = 2.2 for pneumonia). The pre-test probability of pneumonia, derived from published estimates and clinical experience, is estimated to be 20%. Using the graph, the finding of abnormal mental status increases the probability from 20% to 32%; this post-test probability then becomes the pre-test probability for the second finding, diminished breath sounds, which increases probability from 32% to 51%—the overall probability after application of the two findings. Using the approximating rules, both findings (LRs ≈ 2.0) increase the probability about 15%; the post-test probability is thus 20% + 15% + 15% = 50% (an error of only 1%). Using formulas to calculate probability, the LRs of the separate findings are multiplied together and the product is used to convert pre-test into post-test odds. The product of the two LRs is 4.2 (1.9 × 2.2). The pre-test odds would be 0.2/0.8 = 0.25; the post-test odds would be 0.25 × 4.2 = 1.05, which equals a probability of 1.05/2.05 = 51%.

The references for this chapter can be found on www.expertconsult.com.

Please look for the icon throughout the print book, which indicates where the online evidence-based calculator can be used.

REFERENCES

1. Eskelinen M, Ikonen J, Lipponen P. Contributions of history-taking, physical examination, and computer assistance to diagnosis of acute small-bowel obstruction: a prospective study of 1333 patients with acute abdominal pain. *Scand J Gastroenterol*. 1994;29:715–721.
2. Brewer RJ, Golden GT, Hitch DC, Rudolf LE, Wangensteen SL. Abdominal pain: an analysis of 1000 consecutive cases in a university hospital emergency room. *Am J Surg*. 1976;131:219–223.
3. Böhner H, Yang Z, Franke C, Verreet PR, Ohmann C. Simple data from history and physical examination help to exclude bowel obstruction and to avoid radiographic studies in patients with acute abdominal pain. *Eur J Surg*. 1998;164:777–784.
4. Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Ann Emerg Med*. 1992;21(4):384–390.
5. Stiell IG, Greenberg GH, McKnight RD, et al. Decision rules for the use of radiography in acute ankle injuries: refinement and prospective validation. *J Am Med Assoc*. 1993;269:1127–1132.
6. Heckerling PS, Tape TG, Wigton RS, et al. Clinical prediction rule for pulmonary infiltrates. *Ann Intern Med*. 1990;113:664–670.
7. Wells PS, Anderson DR, Bormanis J, et al. Value of assessment of pretest probability of deep-vein thrombosis in clinical management. *Lancet*. 1997;350:1795–1798.
8. Miron MJ, Perrier A, Bounameaux H. Clinical assessment of suspected deep vein thrombosis: comparison between a score and empirical assessment. *J Intern Med*. 2000;247:249–254.
9. Anderson DR, Wells PS, MacLeod B, et al. Thrombosis in the emergency department. *Arch Intern Med*. 1999;159:477–482.
10. Aschwanden M, Labs KH, Jeanneret C, Gehrig A, Jaeger KA. The value of rapid D-dimer testing combined with structured clinical evaluation for the diagnosis of deep vein thrombosis. *J Vasc Surg*. 1999;30:929–935.
11. Funfsinn N, Caliezi C, Baiasiutti FD, et al. Rapid D-dimer testing and pre-test clinical probability in the exclusion of deep venous thrombosis in symptomatic outpatients. *Blood Coagul Fibrinolysis*. 2001;12(3):165–170.
12. Kearon C, Ginsberg JS, Douketis J, et al. Management of suspected deep venous thrombosis in outpatients by using clinical assessment and D-dimer testing. *Ann Intern Med*. 2001;135:108–111.
13. Oudega R, Hoes AW, Moons KGM. The Wells rule does not adequately rule out deep venous thrombosis in primary care patients. *Ann Intern Med*. 2005;143:100–107.
14. Schutgens REG, Ackermark P, Haas FJLM, et al. Combination of a normal D-dimer concentration and a non-high pretest clinical probability score is a safe strategy to exclude deep venous thrombosis. *Circulation*. 2003;107:593–597.
15. Tick LW, Ton E, van Voorthuizen T, et al. Practical diagnostic management of patients with clinically suspected deep vein thrombosis by clinical probability test, compression ultrasonography, and D-dimer test. *Am J Med*. 2002;113:630–635.
16. Chagnon I, Bounameaux H, Aujesky D, et al. Comparison of two clinical prediction rules and implicit assessment among patients with suspected pulmonary embolism. *Am J Med*. 2002;113:269–275.
17. Miniati M, Bottai M, Monti S. Comparison of 3 clinical models for predicting the probability of pulmonary embolism. *Medicine*. 2005;84:107–114.
18. Wells PS, Anderson DR, Rodger M, et al. Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and D-dimer. *Ann Intern Med*. 2001;135:98–107.
19. Wolf SJ, McCubbin TR, Feldhaus KM, Faragher JP, Adcock DM. Prospective validation of Wells criteria in the evaluation of patients with suspected pulmonary embolism. *Ann Emerg Med*. 2004;44:503–510.
20. Newman LG, Waller J, Palestro J, et al. Unsuspected osteomyelitis in diabetic foot ulcers: diagnosis and monitoring by leukocyte scanning with indium In 111 oxyquinoline. *J Am Med Assoc*. 1991;266:1246–1251.

21. Fleischer AE, Didyk AA, Woods JB, Burns SE, Wrobel JS, Armstrong DG. Combined clinical and laboratory testing improves diagnostic accuracy for osteomyelitis in the diabetic foot. *J Foot Ankle Surg*. 2009;48(1):39–46.
22. Ertugrul BM, Savk O, Ozturk B, Cobanoglu M, Oncu S, Sakarya S. The diagnosis of diabetic foot osteomyelitis: examination findings and laboratory values. *Med Sci Monit*. 2009;15(6):CR307–CR312.
23. Rahko PS. Prevalence of regurgitant murmurs in patients with valvular regurgitation detected by Doppler echocardiography. *Ann Intern Med*. 1989;111:466–472.
24. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM*. New York, NY: Churchill Livingstone; 1997.
25. Sackett DL, Haynes RB, Tugwell P. *Clinical Epidemiology: A Basic Science for Clinical Medicine*. 1st ed. Boston, MA: Little, Brown and Co; 1985.
26. Fagan TJ. Nomogram for Bayes' theorem. *N Engl J Med*. 1975;293:257.
27. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. *Lancet*. 2005;365:1500–1505.
28. McGee S. Simplifying likelihood ratios. *J Gen Intern Med*. 2002;17:646–649.
29. Bohadana AB, Peslin R, Uffholtz H. Breath sounds in the clinical assessment of airflow obstruction. *Thorax*. 1978;33:345–351.
30. Pardee NE, Martin CJ, Morgan EH. A test of the practical value of estimating breath sound intensity: breath sounds related to measured ventilatory function. *Chest*. 1976;70(3):341–344.